

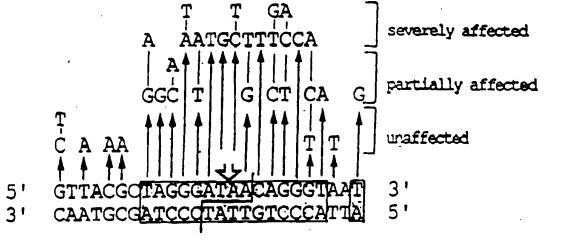
The Universal Code Equivalent of the Mitochandrial 1-See I Gene.

### The synthetic I-Scal gene

Borns HI COGGATOCATE CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GET CCG AAC TOT H H H K N I K K N Q V H N L AAA CTE CTE AAA GAA TAC AAA TCC CAG UTU ATC GAA CTE AAC ATC GAA CAG TTC GAA GCA GET ATC GET CTG ATC CTG GET GAT GET TAC ATC CET TET CET GAT GAA GET AAA ACC TAC 1.  $G \setminus O \setminus A$ TOT ATG CAG TTC GAG TOR AAA AAC AAA GCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT CAG TEG STA CTO TCC CCG CCC CAC AAA AAA GAA COT STT AAC CAC CTS SST AAC CTS STA ATC ACC THE HOC CAR ACT TTC AAA CAC CAA GCT TTC AAC AAA CIE GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA ACC ATC CCR AAC AAC CTE ETT GAA AAC TAC CTR ACC CCG ATG TET CTO BEA TAC TOO TTE ATE GAT GET GET AAA TOO BAT TAC AAC AAA AAC TET ACE 2. n (o o 8 AAC AAA TOR ATO GTA CTG AAC ACC CAR TOT TTC ACT TTC GAA GAA GTA GAA TAC CTG GTT AND GET CTG CET AND AMA TTC CAM CTG MAC TET TAC CTA MAM ATC MAC AMA MAC AMA CCG ATC ATC TAC ATC BAT TOT ATG TOT TAC CTB ATC TAC AAC CTG ATC AAA CCR TAC CTG ATC CCG CAG ATG ATG TAC AAA CTG CCG AAC ACT ATC TCC TCU GAA ACT TIC CTG AAA TAA TAAGTCGACTGCAGGATCCGGTAAGTAAGTAA Sail Patl Bank!

1 and 2: These amino acids are absolutely necessary to produce catalytic activity. Other substitutions are possible, such as deletions of the 10 first amino acids.





1747 TTATASTCCTGTCGCCTTTCCCCACCTCCTGACTTCACCTCCATTTTTCTC ATG CTC AGG SAG SAG SAG 1419 CET ATE COA ANA COC CAG CAA COC COC CET TET ACT CET COT COC CET THE COC C 1879 THE REA CAT GIT CIT TOO TOO GIT ATC CIC TOA TICTOTOGATMACCUTA-METERCETERATIONED 134 28 C 8 H V L 2006 COC CAS TOR OCCUPANCEMENTA ATG THE GITACTICACTICATIACCOMICCONTICACTITI ATG STAR OFF CCC CCC CCT ARC TTC TCT DEA AFT STE AGC CCA TAA CAATITCACACACCAAACACT ATC 2229 ACC ATG AFF BOT ARE TOT CAT CTT TOA CAMCTTATCATCGATAGGTTTA AND CEG TAG TITATCAC 1195 2296 AGTIMATICCIAACGCACICAGGCACIGTGT ATG AAR TCT AMC AAT GGG CTC ATC STC ATC CTC GGC 2363 2364 ACC GTC ACC CTG GAT GCT GTA GGC ASA GGC TTG GTT ASC CGG GTA CTC CCC GGC CTC TTG 2423 2424 COC GAZ ATC COC CTG ATG COT GAA COT GAC GGA COT AAC CAC CAC ATG TOT CTC CTC 2433 2484 THE USE TOO OCA TOO CAG CAC AME THE TOO TOO GOT AME CTG CTG AGE COG OCH AME CTT 2542 2944 ACT CCC CAT LUL CET GIT CAC AAT TAA TCATCGCCTCGTATA ATG TGT CCA ATT GTG ACC CGA 7404 Berry 2007 THE CHAPTERCHOLOGORANCE SCREEN ATE LAT ATTS AND AND ATE AND AND CAG GTA ATE 2612 2871 AME CTC COT CCC AME TOT AMA CTG CTG AMA GAS THE AMA TOC CMG CTG ATC CAS CTG AMC 2730 LRBYES 2731 AFC GAA CAG PRY GAA GCA CGF AFC GCF CFC AFC CF0 667 GAT 6CT TAC AFC CGT JUT GGT A G I G L I L G D A Y I 2791 GAT GAA GGT AAA ACC THE TOT ATG CAG TTC CAG TCC AAA AAC AAA GCA THE ATG CAC CAC 53 0 E G R T Y C H O F E H R H R A Y H O R 2851 GEA TOT CTG CTG TAC CAT CAG TIME ITTE CTG TCC CCG CCC ANA ANA CAA COT OTT AMC 2910 Q W V L S 2971 ANA CTG CCT AMC CTG TEC ATC GTT AMC AMC AMA ACC ATC CCG AMC AMC CTG GTE CAM. 1031 AMC THE LITE ACC COS ARE TOT CTG OCK THE TOG TTC ATG GAT GAF OCT GAT ANA TOG CAF 3090 3091 THE MAR MAR AND THE MAR AND TEG ATE GTA CTG AND ACC CAM TET TTC ACT TTC CAA 3150 SIEL CAA CEA CAA PAC CES CET AMA GOT CES CET AME AMA THE CAA CEG AME TOT INC CEA AMA 173 E V B C Y V E 1210 1 92 1911 ATC AMC MAN AMC AMA COM ATC ATC TAG AZC GAZ TOT ATG TOT TAG CTG ATC TTC TAG AMC 183 T L . F T H 3270 3271 CTG ASC AAA CCG TMC CTC ASC CCC CAG AGG ATG TMC AAA CTG CCC AAC ACT ATU TUU TUU 3330 213 T. T E P Y L I P Q N N T R L P E I I S S 232 1331 CAN ACT THE CITE AND THE TANGETCENCTOCKECTTOCKACTOCCCCCCTTTLACACCTCCCCCCC

I-SceI coding sequence of pSCM525 - Note the two amino acid N-terminal extension-as compared to genuine version of the gene.

### VARIATIONS AROUND THE 1-SCR ! SECUENCE

		-2 H_	•1 	1	_ <b>_</b>	¥	_ـــــــــــــــــــــــــــــــــــــ	5	<u> </u>	<u> </u>	_1	¥	10	4	L	:	þ	4	š
K	. !	<u>_</u>	20 K	٤	4	K	\$	a	L	I	ξ	Ŀ	30 N	ı	٤	:	:	ï	7
G	1	£	40	T	<u>L</u>	ā	2	A	I	1	4	s	<del>ک</del> ن ۲	2	٤	3	<	•	1
c	Ħ	•	60 F	ε	W	ĸ	н	ĸ	<b>A</b>	Y	н	0	70 H	٧	ς	L	-	*	:
ą.	٧	Ą	80	3	P	P	н	K	K	٤	R	¥	90 N	Ħ	L	4	Ħ	-	•
1	T	v	100 G	4	٥	T	F	K	H	٥	A	F	110 N	ĸ	L	À	N	-	:
1	٧	N	120 N	K.	K	I	ı	P	N	N	L		130 E		*	<b>L</b>	1	2	•
3	L	A	140 Y		٢	ក	2	0	2	3	*		150		H	K	N	s	:
N	ĸ	s	160 1	٧	L	N	T	Q	3	F	t	F	170 E	٤	Ą	E	Y	-	¥
K	e	L	180 R	N	K	F	a	L	N	c	¥	٧			N	K	N	•	>
1	1	Y	200	<b>'</b>	6	ង	s	Y	L	ı	F	Y			t	K	P	Y	-
1	P	q	220 H	я	¥	K	L	P	N	T	ī	3	230	٤	T	F	L	<	•

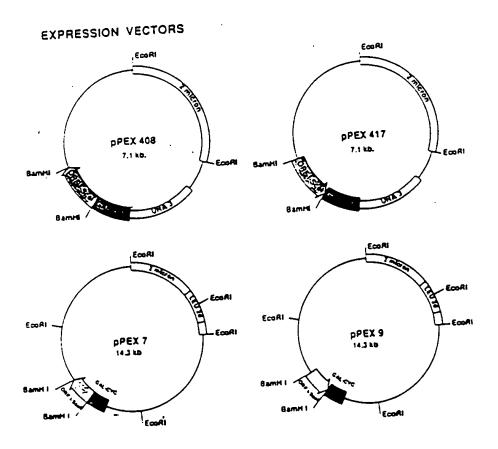
Positions then can be changed without affecting ensume activity (demonstrated)
positions -1 and -3 are not natural. The two amine acids are added due to cloning strategies
position 36: Q is tolerated
position 40: M or Y are tolerated
position 41: S or M are tolerated
position 43: A is tolerated
position 43: A is tolerated
position 45: Y or N are tolerated
position 91: A is tolerated
position 133 and 153: L are tolerated
position 133: A and S are tolerated

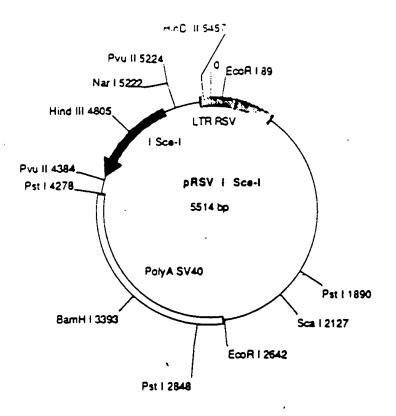
Changes that affect ensume activity (domenstrated) position 19: L to A position 39: I to A or M. position 49: L to A position 40: L to A position 44: L to A position 44: L to A position 44: A to E or H position 45: A to E or D position 45: Y to D position 47: I to A or M position 47: I to B or M position 47: I to B or M position 47: I to B or M position 14:D to E position 145:D to E position 145:D to E position 140:Q to E

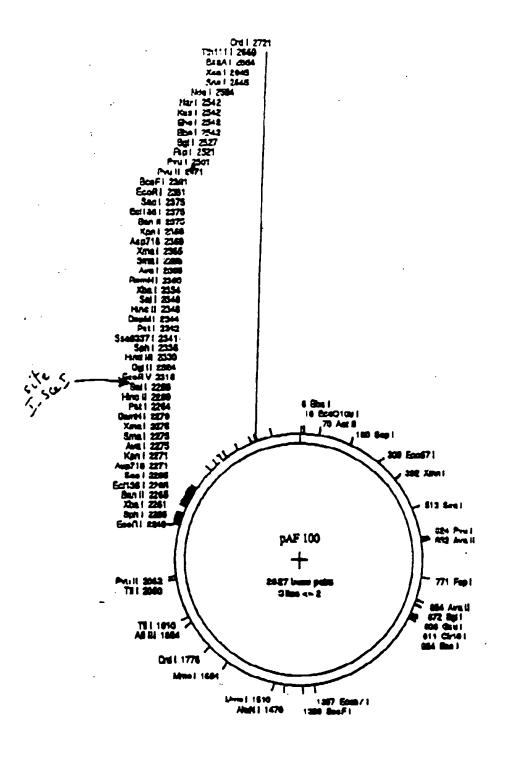


### Group I Intron Encoded Endonucleases and Related Endonucleases

	Endonuclease	Recognition Sequence	CUPAYA CE	▼Intron site
	I-6cs I (Seotheromyces mitochondria)	CGCTAGGATAI	S 2 C C 2 A	ATATASE TATASES
	I-See IV (Seccheromyces mitochondria)	TTCTCATCATTA		
7	l-See il (Seccharomyces mitochondria)	GAAACCAGTAGG	ASAAS!A	*
ante	I-Cau I (Chlentydomoues chloropiest)	TARCOGICE CAN	FCCATCGC	TTTAAST
d 1 July	i-Ppo i (Physarum nucleus)	t GAC CC CC CCA	reentess W	****
Two Dodecapaide Family for 4 hp cutteral	I-Sos III (Saccheromyces mitochondria)	CCTCCRAAACC	TARIPANTA ARRAPTAA	
abite	(Chlamydomonas chloroplast)		G T G T G A G C A G C A C ° C	4646771
rodeca	Endo See I (RF3) (Seccharomycas mitochondria) (Non is reste)	CTACGACATCC		AT A A D D A D A A D A A D A A D A A D A A D A A D A A D A A D A A D A A D A A D A
Teal	HO (Saccharomycan mucleus) (Non inworld)	CTTTCCGCAAC	1 4 4 7 A 7 P A 7	111414
	[-Cam I (Chlamydomones mitochondria) (Puntos endomones)	A C C A T G G G G T C	A A T U T U T T T T A C A G A	TTCTGGG AAGACCC
-	I-Pan I (Podospeara mitochondria) (Podospeara mitochondria)	G 1 G C C T G A A F G C A C C C A C T T A C	1	TREETTT .
Families	(Becomlophage T4)		<b>y</b>	
Te and	diffect A	O TA GAT O TTTTCTT UA TUTA CAAAAAAA	CCCAGATG	acalitta
Ouchural	I-Tev II C A A G I I	GCTTATGAETATGA CJAATACTCATACT	AGTGAACA TCACTTGT	66 7 7 7 7 7
Other Bir	PTer III GCTATTCG	rrittatetatett Liliforatidal	* * 6 6 6 7 6 *	1661111





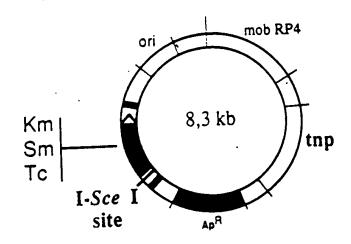


```
PAF 100 - RESTRICTION MAP
                                                     Sau JA I
                                                     HOO I
                                                 Serr I
                                                 Rel I
                                                 Mop I
Ipa II
                                                 Dan V
                                                 315% ]
                                                 SHA I
SHA I
                                                 Scrf
                                                 NCL I
                                                    NIa IV
                                             Kpn I
                                                 Batk I
                                      Halk
                                      Ec1136 :
                                                 BaaJ
                                       Bapiles I Ben I
                                                 Ben I
                                       Ban II
                           3507524 I Alu I
Nap I Rea I A
                                                 AVE 1
                                             Asp718 ALV 1
             CCAMCTOCAATTCCCATCCTCTAGACCTCCCTACCCGGGATCCTCCAGTCCACCCCACTACACGATACACGATACACACT
             SCHICCARCTTARCOTTACGACATCTCCARCCATCCCCCCTACGACGTCACCTCCCTATTGTCCCATAAAGTCTA
                                            *11 [1 [1 ]
2271 2274
                                 -11 11
                                    2262
                1 | 1 | 2244
                                                                 2286
                                             2271
                            2255
2255
                                                  2275
1 2279
2275
2276
                                                                  2292
                   2247
                                   2261 2265
                                              2271
                      2249
                            2255
                            2255
                                       7245
                                                                   2290
                                                   2275
                                       2265
2860
2265
                                                  2273
                                                       1280
                                               2272
                                                       2279
                                              2271
2271
                                                        228Q
                                                2277
                                                   2275
                                                   2275
                                                   2275
                                                   2275
                                                    2275
                                                     2276
                                                     2274
                                                     2276
                                                     2276
                                                     2276
                                                         2280
2280
                                                          2280
```

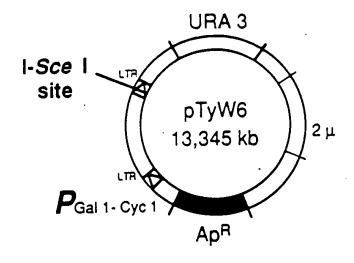
```
paf 100 - RESTRICTION MAP
                                                                                                                                             and :
Acr :
acil :
                                                                                                                                              5P4 1:
                                                                                                                                             324 V
                                                                                                                                              SALK I
                                                                                                                                             Ban I
                                                                                                                                           XEA [
                                                                                                                                           Serf I
                                                                                                                                           Mer I
                                                                                                                                           014 V
                                                                                                                                           RATE T
                                                                                                                                           BaaJ !
                                                                                                                                           Ben !
                                                                                                                                 I KEPPS
                                                                                                                                                                     Sac T
                                                                                                                                 Hoo I
                                                                                                                                                                     Rgia I
                                                                                                                                Opn II
Opn I
                                                                                                                                                                     Ec1116 :
                                                 Ala I
                                                                              Sia :
Pat :
                                                                                                      Ple I
                                                                                                                                                        RMA I
                                                                                                      I thiE
                                                                                                                                                        Capé I
                                                                                                                              NLA IV
                                          Dde I
                                                                            See8 117
                                                                                                                              BatY I
                                                                                                                                                     AIT LA
                                  SAGJA I
                                                                                                 Iaq I
                                                                                                                              Band I
                                                                                                                                                   Kpa :
                                                                 NIA III
                                                                                                                                                                                    EcoR :
                                  I. odes
                                                                                                                                                     Bea I Alu I
                                                                                                                                                                                                         Sdl ::
                                                                                                                              Alw I
                                                                                              SA1 I
                                  דד שלה
                                                               Spn I
                                                                                              Hinc II
                                                                                                                   HAL I AVE I
                                                                                                                                                                     3sp1286 [
                                                              BEPC I
                                  Opa I
                                                              Esp7524 I
                                                                                          ACC I SEA I
                                                                                                                                   Meau I
                                                                                                                                                                     San II
                                                                                                                                                                                                          1 943
                                BULT I
                                                                                                             XDa I Alv L Asp/LH Tag I
                                                              Hap I AspH.I
                                                                                                                                                                                                    Bar I
                                                                                                                                                                                                    1 11:
                                                               11
                                                                          11.1
                                                                                            . . . .
                        APCAMATCIANDETTOCATOCCTGCADDICCACTCTAGAGGATCCCGGGGTACLUALUTUMACTCACTGGCCGTCCTTT 2400
                        THE TOTAL COTTACO TACO CALCOTOCO ACTO A CALCOTOCO ACTOCO CONTROL CONTROL CALCOTA ACTOCA ACTOC
                                                               1.14
                                11 1 (1)
                                                                                                              2354
                                                                                                                              2261
                                                                                                                                                                                                    2367
                                                               2336
                                                                                   2311
                                2324
                                                                                                                        58 2364
2388 236
                                                                                             22(4
                                                                                                               :358
                                                                                                                                                                     2375
                                                                                                                                                                                                        2389
                                2324
                                                               2236
                                                                                                                                                                     2375
                                                                                                                                        2265
                                                                                                                                                                                                          2349
                                                               2326
                                                                                              2340
                                  2325
                                                                                                                                                     2369
                                                                                              2348
                                                                                                                              2360
                                                                                                                                                                     2376
                                                                                                                                                                                                            2300
                                                               2136
                                  2325
                                                                                                                              2160
                                                                                                                                                     2369
                                                                                                                                                                                    2341
                                                                                                                                                                                                              2391
                                                                 2337
                                                                                                 2349
                                  2325
                                                                                                                              2160
                                                                                                                                                     2368
                                                                            2341
                                  2325
2351
                                                                                                                              2760
                                                                                                                                                        2370
                                          2328
                                                                              2342
                                                                                                      7351
                                                                                                                                 2361
                                                                                                                                                        2370
                                                                              2342
                                               2330
                                                                                                                                 2361
                                                                                                                                                                     2375
                                                  2331
                                                                                                                                                                     2375
                                                                                                                                 2361
                                                                                                                                                                     2375
                                                                                                                                  2361
                                                                                                                                           2343
                                                                                                                                           2365
                                                                                                                                            2365
                                                                                                                                            2365
                                                                                                                                            236
                                                                                                                                            2365
                                                                                                                                              2366
                                                                                                                                              2366
                                                                                                                                              2366
                                                                                                                                               2366
                                                                                                                                              2344
                                                                                                                                              2366
                                                                                                    Mae III
                                                                                     Sarf I
                                                                                      Real II
                                                                                     004 V
                                                                                                                                                                                                               Pos II
                                                                                                                                                            Pag4E
                                                                                      BetT T
                                                                                    COLTGOCCTTLACECAACTEAATCCCCCTTCCLACCACACCCCTCTVTTCCCCCAACCTCCCC
                                                                                            COCCALIGOETTEANTINGCOCAN
                                                 1
                                                                                                                                                                                                               2471
                                                                                                                                2440
                                                                                                                                                            2451
                                     2405
                                                         2413
                                                                                    2423
                                                                                                                                                        . 2451
                                                                                                                                                                                                                2471
                                                                                      2424
                                                  2410
                                                                                                                                                                                                                  2472
                                                                                                                                                                           2457
                                                                                      2424
2424
                                                                                      2424
                                                                                      2424
```

2430

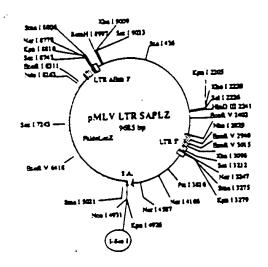
Names: pTSm ω pTKm ω pTTc ω

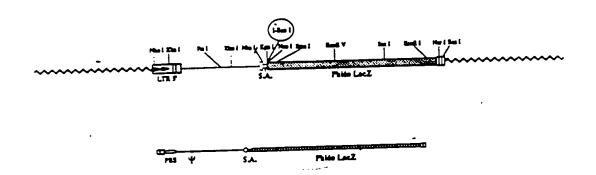


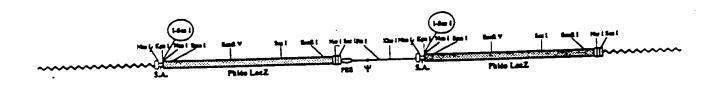
Construction: pGP 704 from De Lorenzo, with transposase gene and insertion of the linker [I-Scel] in Notl unique site

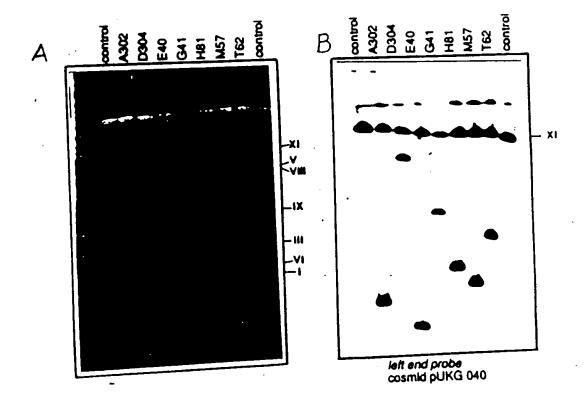


Construction: pD 123, from J. D. Boeke with insertion of a linker [I-Scel-Notl] in BamHI



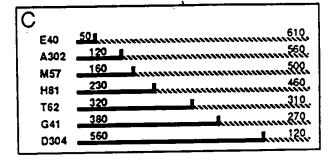


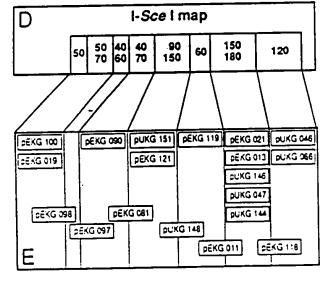


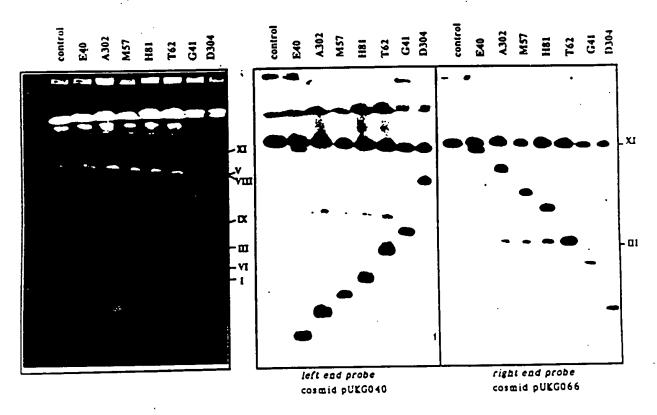


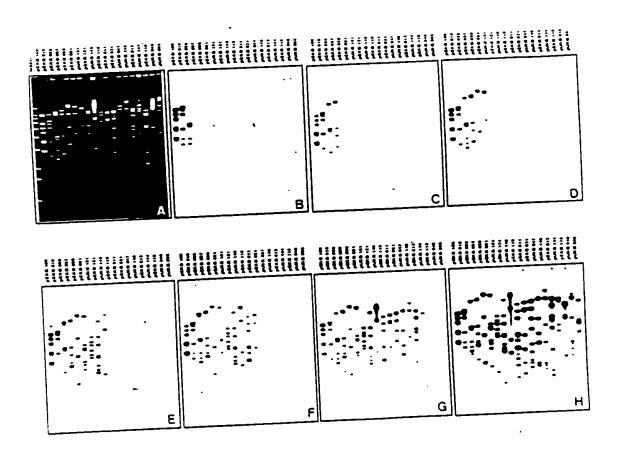
Α	
Δ302	Brown William William William William
E40	.50
G41	270
H81	
M57	
T62	310

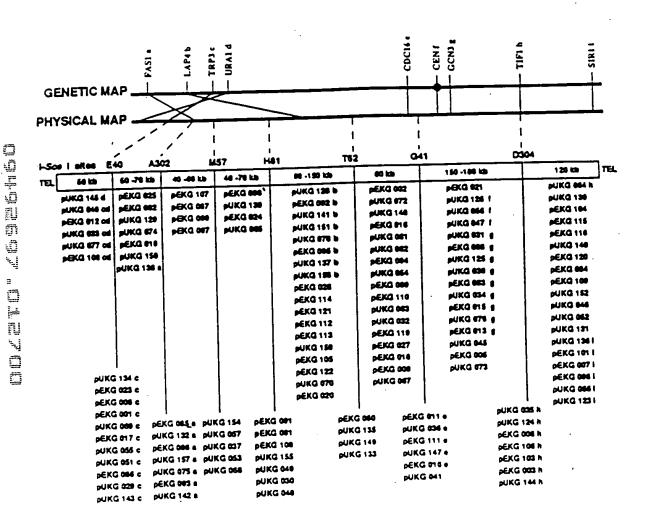
B A302	120 <b>1</b>	560
E40 G41	380	
H81	230	
M57	160	500
T62	320	31 <u>0</u>

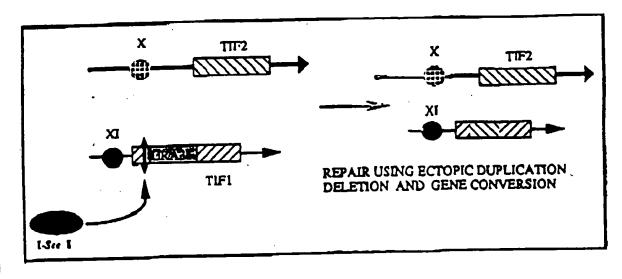


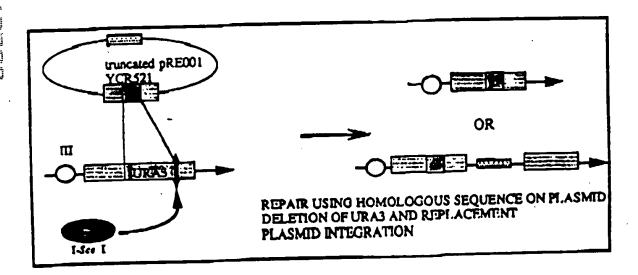




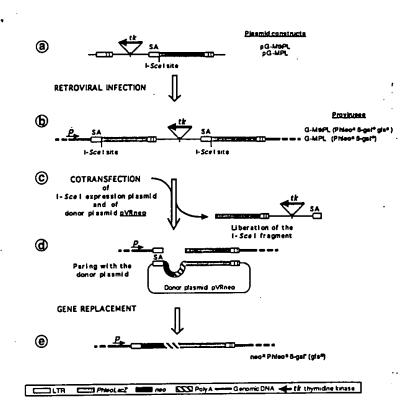








### Figure 20



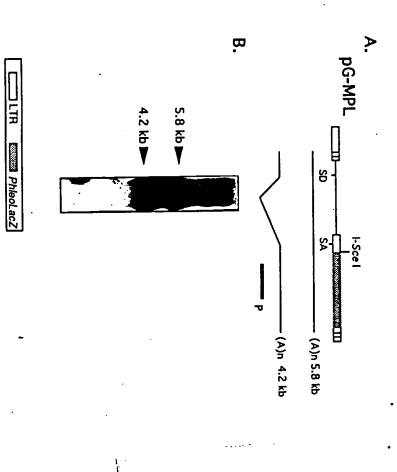
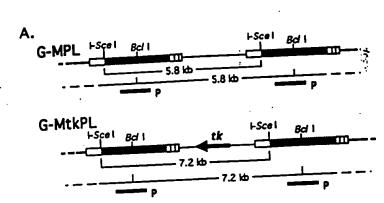


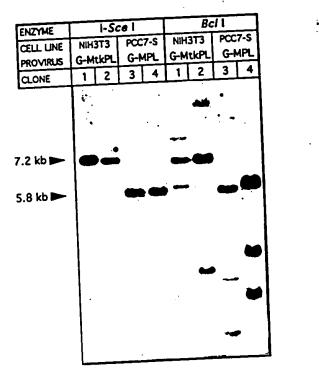
Figure 21

*(*. , .



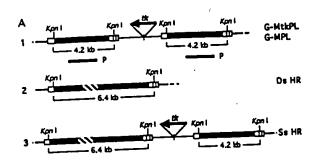
В.

;



LTR PhleoLacZ +tk thymidine kinase

В



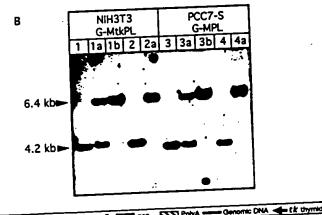
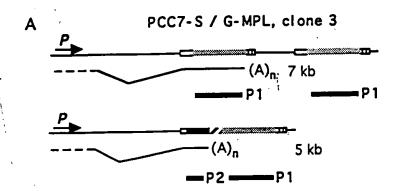
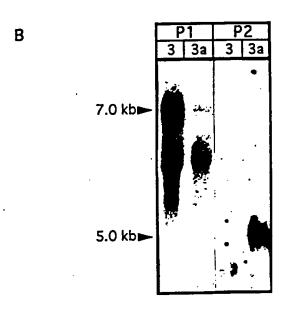
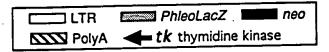


Figure 24







### Figure 25

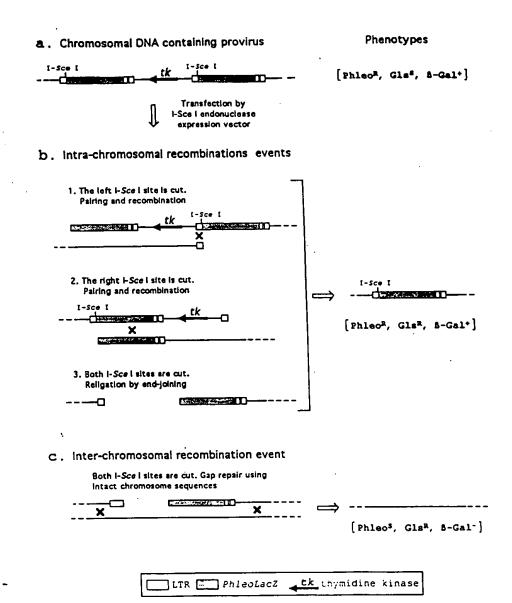
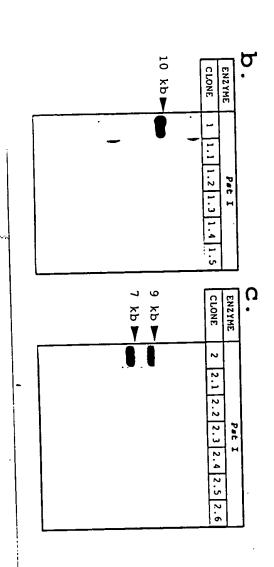


Figure 26

a. Parental DNA, G-MtkPL





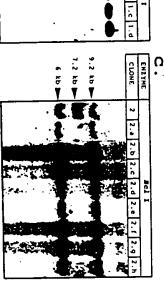
odanasta cometa

## 1. Parental DNA, G-MtkPL



# 2. Intra-molecular recombination event



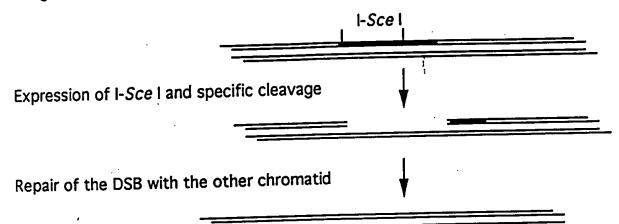


11.6 Kb T

DS49252.C12700

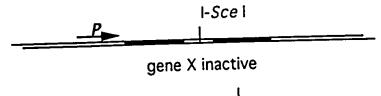
### LOSS OF HETEROZYGOSITY

Integration of artificial site or presence of specific site

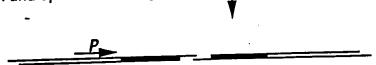


### CONDITIONNAL ACTIVATION (Tandem repeat) Integration of artificial site between tandem repeats

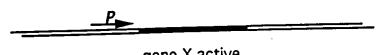
Integration of artificial site between tandem repeats



Expression of I-Sce I and specific cleavage



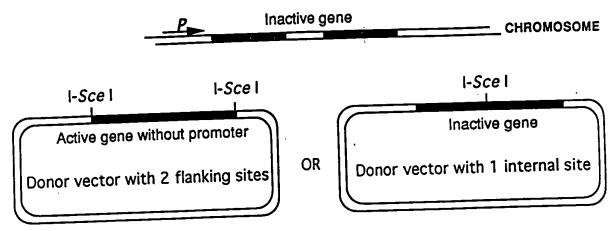
Repair of the DSB by single strand annealing



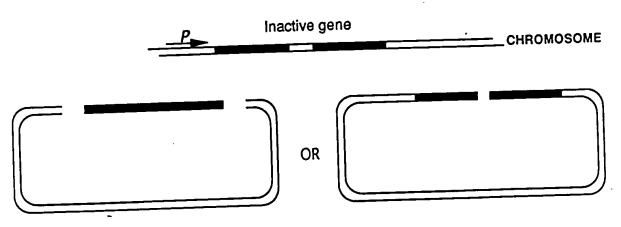
gene X active



### Integration of artificial site or presence of specific site



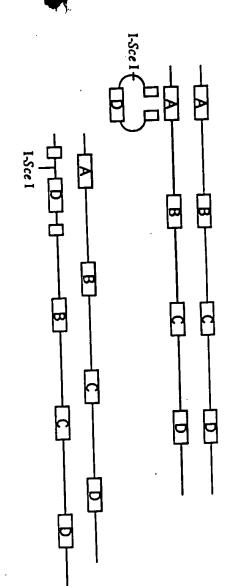
Expression of I-Sce I enzyme and specific cleavage of the donor plasmid



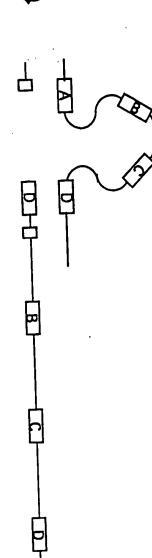
Recombination between the chromosome and plasmid



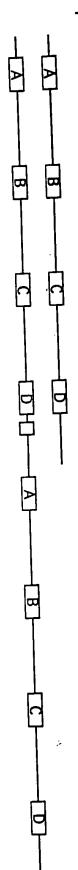
1 insertion of I-Sce I site by classical gene replacement



Specific cleavage by I-Sce I enzyme and repair of the break by homologous sequences

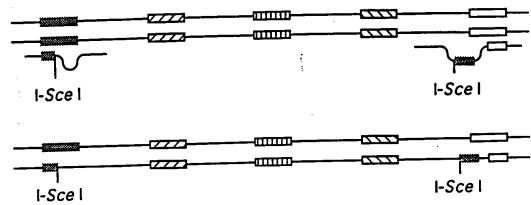


3 Duplication of the totality of the locus

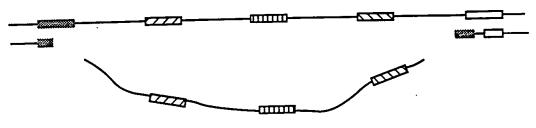


### DELETION OF ACCUS

1 Insertion of two I-Sce I sites flanking the locus



2 Expression of the enzyme and cleavage



3 Recombination between the two ends



4 deletion of the locus

